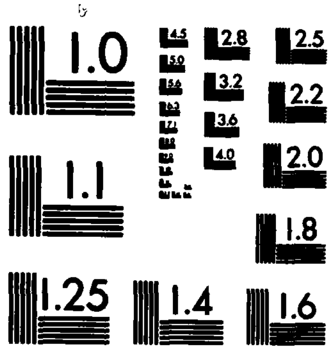
AD-A120 474 NEW TECHNIQUES FOR MEASURING SINGLE EVENT RELATED BRAIN POTENTIALS..(U) PURDUE UNIV LAFAYETTE IN SCHOOL OF ELECTRICAL ENGINEERING C D MCGILLEM ET AL. 01 SEP 82 1/1

UNCLASSIFIED AFOSR-TR-82-0901 AFOSR-80-0152 F/G 6/16 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

5

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 82-0901	2. GOVT ACCESSION NO. AD-A120474	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) New Techniques for Measuring Single Event Related Brain Potentials		5. TYPE OF REPORT & PERIOD COVERED Interim Report 1 April 1981 - 30 Mar 1982
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) C.D. McGillem J.I. Aunon		8. CONTRACT OR GRANT NUMBER(s) AFOSR-80-0152
9. PERFORMING ORGANIZATION NAME AND ADDRESS School of Electrical Engineering Purdue University W. Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2313/A4
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NL Bolling AFB DC 20332		12. REPORT DATE 9/1/82
		13. NUMBER OF PAGES 15
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Event related potential; evoked potential, pattern classification; linear discriminant analysis, quadratic discriminant analysis, data preprocessing, visual evoked potential.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Alternative methods of selecting features of visual evoked potentials for automatic pattern classifications are compared. Forward sequential feature selection with linear and quadratic discriminant functions, step-wise linear discriminant analysis and exhaustive enumeration with a linear discriminant function are considered. It is found that exhaustive enumeration provides a moderate improvement over the other procedures. In many cases the optimum set of features selected for a given size set does not contain the same features

20. Abstract (cont.)

that were found optimum in a set having a different number of features.

A new class of filters has been developed for processing evoked potentials that provides much greater discrimination against the ongoing EEG than other filters that have been employed. The filter incorporates all deterministic and statistical information about the signal and noise into a matrix operator that processes the measured data vector. Typical results are shown.

A study of the effects of the ongoing EEG on measurement of the latency of evoked potential components is described. The results of the study indicate that the majority of the latency variations found experimentally are intrinsic to the evoked potentials themselves and do not result from interference by the ongoing EEG.

RESEARCH OBJECTIVES

The primary research objective of this project is to develop and evaluate methods of measuring the parameters of single evoked potential (EP) waveforms. Currently the research falls into the following two areas.

a. Investigation of pattern recognition procedures for discriminating among various evoked potential waveforms.

b. Investigation of preprocessing and filtering techniques to provide improved waveform estimation.

STATUS OF RESEARCH EFFORTS

Significant results have been obtained in the areas of research relating to this project. These are briefly described in the following paragraphs and are further described in the publications resulting from the research.

Feature Selection for Automatic Classification. It has been shown by simulation and experiment that it is possible to effectively classify EP waveforms into categories corresponding to the stimuli that produced them. Two problems associated with this procedure are selection of the best features from the data set to use in making the classification and the errors that result from interference due to the ongoing EEG. Both of these problems have been addressed as part of this research project and procedures developed to mitigate their effects on system performance.

In the classification procedures being utilized here the features upon which a decision is made are the waveform amplitudes at regularly spaced sampling intervals. These samples are taken from a segment of the EP waveform immediately following stimulation and extending up to as much as 500 ms. The sampling interval is typically 4 to 20 ms. It is a general property of classification procedures of this type that increasing the number of features employed in the classification improves performance up to a certain (relatively small) number of features after which performance begins to decrease. The problem is to select the best subset of features from the complete set.

Two methods that have been widely used for feature selection are Forward Sequential Feature Selection (FSFS) and Stepwise Linear Discriminant Analysis (SLDA). The FSFS algorithm performs a classification of the training set using each feature individually in a sequential manner. The feature yielding the lowest error rate of classification is then selected. The next step is to combine this first feature with each of the other features and reclassify the training set to pick the pair of features that gives the lowest error rate. This process is continued until no further improvements in accuracy are obtained by adding additional features or until the desired number of features has been selected. This is not an optimum procedure because there is no basis for assuming that because one feature works the best alone that that feature paired with any other feature would be the best two features for classification and the same thing applies for larger numbers of features.

The SLDA algorithm proceeds in much the same manner as the FSFS algorithm except that the criterion used to select the best feature at each step is based upon a statistical test rather than a computed error rate. The test used is a one way analysis of variance. The SLDA also tests for loss of significance of any of the features already entered and thus can remove variables or features that have been previously selected. However, in the applications that we have made of this procedure, previously selected variables have not been removed by the procedure. This procedure is also a nonoptimum procedure and is only applicable to linear discriminant analysis.

One method of obtaining an optimum feature subset is to test all possible subsets for their performance in a classifier. The difficulty with this procedure is that a great amount of computation is required in order to test every combination of features that is available from a large set. For example, selecting five features out of a set of 27 would require the design and testing of 100,000 different classifiers. However, it was found that by writing all computer subroutines in assembly language and by calculating the discriminate function using the algorithm for regression by "leaps and bounds"* the computation could be made feasible for linear discriminant analysis. Using this procedure it is then possible to compute the optimum feature set and to compare performance with the optimum feature set with those features selected by the FSFS and SLDA algorithms. Also some additional information can be obtained by examining which features are selected by each procedure to see what commonality exists among them.

The classification performance using features selected by each of the three techniques was measured using both artificial simulation data and using real evoked potential data. The simulated data was generated by adding known deterministic signals to random noise sequences. The noise was generated as a Markov process having various half-power bandwidths. The signals corresponded to average evoked potential waveforms measured experimentally using the unexpected event paradigm in which a letter regularly appears but is inverted in a random manner approximately 10% of the time. An experiment of this type is the source of the real data that is considered subsequently. Tests were made using noise of various bandwidths ranging from 4 to 25 hertz and evaluation was made by training and testing on the same data set. Eighty samples in each class were used in the simulation. Three linear classifiers using features selected by FSFS, SLDA, and Exhaustive Search Feature Selection (ESFS) and a quadratic classifier using FSFS were evaluated.

*Furnival, G. M. and Wilson, R. W., "Regression by Leaps and Bounds," *Technometrics*, Vol. 16, 1974.

Table 1 shows the results for noise bandwidths of 8 hertz and for signal-to-noise ratios of -3dB, -6dB and -9dB. For this case the noise bandwidths were the same for the two classes and so their covariance matrices were also identical for the two classes. Table 2 shows the results for the case when the covariance matrices of the noise are different for the two classes as a result of using a noise bandwidth of 8 hertz for one class and 14 hertz for the other class. It is seen from Table 2 that ESFS has the same kind of improvement in performance for the case of unequal covariance matrices as it did for equal covariance matrices for the two classes. The FSFS procedure with a quadratic discriminant function seems to give somewhat better performance for the unequal covariance matrices than do the linear discriminant functions.

A comparison of the several feature selection and classification procedures using measured evoked potentials was also carried out. Data was collected from four human subjects for the unexpected event paradigm. A sequence of letter v's was shown to the subject with a random occurrence of an inverted v occurring 10% of the time. For each subject three electrode sites were used, Cz, Pz and Oz. Feature selections and classifications for all electrode sites were then carried out. Table 3 shows the results for one of the subjects. As in the case of the simulated data the ESFS procedure gave a modest improvement in classification performance over the other linear methods. However, the quadratic discriminant function using FSFS in a number of cases gave better performance than the ESFS procedure.

In addition to the classification results using the experimental data a record was also made of the features actually selected by the different procedures. Table 4 shows the features that were selected at the various steps for the same subject as in Table 3. In most instances it was found that the first feature selected was the same by all methods but from then on there was often divergence particularly with the ESFS algorithm. Oftentimes the initial feature chosen or early feature sets are not contained in the final feature sets. An example of this is shown in Table 4 for electrode Cz, where none of the features selected in Steps 1 and 2 were included in the feature set obtained at Step 5.

The general conclusion of these tests is that the exhaustive search feature selection procedure gives a modest improvement in performance of the classifier for both the simulated and experimental data. In almost all cases features selected at the various steps of ESFS did not contain all the features selected in previous steps. Features were constantly replaced by new ones when moving to the next step. This was never the case for SLDA even though SLDA contained the capability of removing features at each step.

Table 1

Artificial, equal covar. Data sets 5e and 5u, -3dB S/N.

Noise Bandwidth	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
BW1 8hz BW2 8hz	1	72.50%	72.50%	65.00%	72.50%
	2	76.25%	76.25%	73.75%	78.75%
	3	81.25%	81.25%	80.00%	83.75%
	4	83.75%	83.75%	86.25%	88.75%
	5	83.75%	83.75%	85.00%	91.25%

Artificial, equal covar. Data sets 5e and 5u, -6dB S/N.

Noise Bandwidth	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
BW1 8hz BW2 8hz	1	70.00%	70.00%	60.00%	70.00%
	2	71.25%	71.25%	67.50%	73.75%
	3	75.00%	75.00%	70.00%	77.50%
	4	76.25%	76.25%	77.50%	81.25%
	5	77.50%	77.50%	76.25%	83.75%

Artificial, equal covar. Data sets 5e and 5u. -9dB S/N.

Noise Bandwidth	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
BW1 8hz BW2 8hz	1	63.75%	63.75%	56.25%	63.75%
	2	68.75%	68.75%	60.00%	68.75%
	3	70.00%	70.00%	66.25%	72.50%
	4	71.25%	71.25%	66.75%	75.00%
	5	72.50%	72.50%	68.75%	76.25%

Table 2

. Artificial, unequal covar. Data sets 2e and 6u, -3dB S/N.

Noise Bandwidth	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
BW1 6hz BW2 14hz	1	71.25%	71.25%	70.00%	71.25%
	2	77.50%	77.50%	77.50%	78.75%
	3	82.50%	82.50%	78.75%	83.75%
	4	87.50%	87.50%	82.50%	87.50%
	5	88.75%	87.50%	86.25%	88.75%

Artificial, unequal covar. Data sets 2e and 6u, -6dB S/N.

Noise Bandwidth	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
BW1 8hz BW2 14hz	1	68.75%	68.75%	66.25%	68.75%
	2	72.50%	72.50%	70.00%	72.50%
	3	76.25%	75.00%	73.75%	77.50%
	4	78.75%	76.25%	73.75%	80.00%
	5	78.75%	76.25%	76.25%	85.00%

Artificial, unequal covar. Data sets 2e and 6u, -9dB S/N.

Noise Bandwidth	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
BW1 8hz BW2 14hz	1	67.50%	67.50%	60.00%	67.50%
	2	67.50%	70.00%	65.00%	68.75%
	3	67.50%	71.25%	70.00%	71.25%
	4	70.00%	72.50%	66.25%	75.00%
	5	71.25%	71.25%	71.25%	76.25%

Table 3

Real data, error rate. Subject 3. Electrode Oz.

Data Structure	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
expected–unexpected event	1	75.00%	78.75%	75.00%	75.00%
	2	82.50%	86.25%	82.50%	82.50%
	3	87.50%	87.50%	83.75%	87.50%
	4	87.50%	87.50%	85.00%	87.50%
	5	88.75%	87.50%	85.00%	88.75%

Real data, error rate. Subject 3. Electrode Pz.

Data Structure	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
expected–unexpected event	1	76.25%	75.00%	76.25%	76.25%
	2	83.75%	83.75%	83.75%	83.75%
	3	86.25%	86.25%	82.50%	87.50%
	4	86.25%	90.00%	83.75%	88.75%
	5	86.25%	90.00%	86.25%	91.25%

Real data, error rate. Subject 3. Electrode Cz.

Data Structure	Step Number	Average Classification Rate			
		FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
expected–unexpected event	1	73.75%	73.75%	72.50%	73.75%
	2	85.00%	83.75%	80.00%	85.00%
	3	86.25%	86.25%	82.50%	86.25%
	4	86.25%	87.50%	83.75%	87.50%
	5	86.25%	88.75%	82.50%	90.00%

Table 4

Real data. Subject 3. Electrode Oz.

Step	Latency of Feature(s) Selected (milliseconds)			
	FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
1	576	576	576	576
2	576 196	576 196	576 196	196 576
3	576 196 216	576 196 176	576 196 276	196 216 576
4	576 196 216 136	576 196 176 236	576 196 276 476	136 196 216 576
5	576 196 216 136 256	576 196 176 236 436	576 196 276 476 336	136 176 196 216 576

Real data. Subject 3. Electrode Pz.

Step	Latency of Feature(s) Selected (milliseconds)			
	FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
1	576	556	576	576
2	576 296	556 296	576 296	296 576
3	576 296 176	556 296 456	576 296 476	296 476 556
4	576 296 176 236	556 296 456 476	576 296 476 156	156 316 336 476
5	576 296 176 236 256	556 296 456 476 196	576 296 476 156 376	156 316 396 476 656

Real data. Subject 3. Electrode Cz.

Step	Latency of Feature(s) Selected (milliseconds)			
	FSFS(linear)	FSFS(quadratic)	SLDA	ESFS
1	556	296	576	556
2	556 296	296 556	576 296	296 556
3	556 296 136	296 556 156	576 296 476	136 296 556
4	556 296 136 196	296 556 156 396	576 296 476 376	156 296 456 556
5	556 296 136 196 216	296 556 156 396 276	576 296 476 376 656	136 296 316 456 576

Time-Varying Filter In the measurement and analysis of evoked potentials it is generally known within what time epoch the signal occurred. However, often the signal can take a variety of forms through variations in latency amplitude or even occurrence of particular components in the EP. By incorporating all known apriori information both deterministic and probabilistic into the processor design it is possible to obtain substantial improvements in waveform estimation over processors not utilizing this information. Because of the signal's occurrence in a particular time epoch in combination with the ongoing EEG the resulting waveform is a sample function of a non-stationary random process and the optimum processor takes the form of a time-varying filter. For processing sampled data the filter is a matrix operator whose elements are determined from the known or assumed parameters of the underlying random processes associated with the signal and the ongoing EEG. The filter operator is selected to minimize the mean square error of the estimate and is given by the following expression

$$\underline{H} = \underline{K}_{sx}\underline{K}_{xx}^{-1}$$

where $\underline{K}_{sx}$ is the cross covariance matrix of the signal and data and $\underline{K}_{xx}$ is the covariance matrix of the measured data.

The matrix $\underline{H}$ can be thought of as an operator that projects the data vector from the high dimensionality measurement space into the low dimensionality signal space. The dimensionality of the signal s^ ce is determined by the number of significant eigenvectors of the operator and this in turn depends upon the covariance matrices themselves. To a first approximatior the more that is known about a signal the lower will be the dimensionality of the signal space. For example if the signal were deterministic except for amplitude the signal space would be one dimensional and only the signal waveform would appear at the filter output regardless of the input.

The matrix operator filter design can be carried out for EP waveforms in the following manner. From an ensemble of measured waveforms the latency corrected average (LCA) is computed. From the output of the LCA procedure the individual components in the waveform are identified, their shapes determined and the means and standard deviations of their latency variations estimated. From this information the covariance matrix of the signal can be determined and the covariance matrix of the measured data is calculated directly from the data set. From these two matrices the filter matrix is computed.

Filters have been designed and tested using both simulated and measured data. Figure 1 shows an example of filtering simulated data. Figure 1a is the signal, Figure 1b shows samples of signal plus noise and Figure 1c shows the filtered versions of Figure 1b. A filter was designed for EP waveforms obtained using a checkerboard visual stimulus. Figure 2 shows individual measured EP waveforms (dashed) and filtered versions of the waveforms (solid). The very substantial noise reduction performance of this filter is readily evident in igures ' .ad 2. The performance of this type of filter

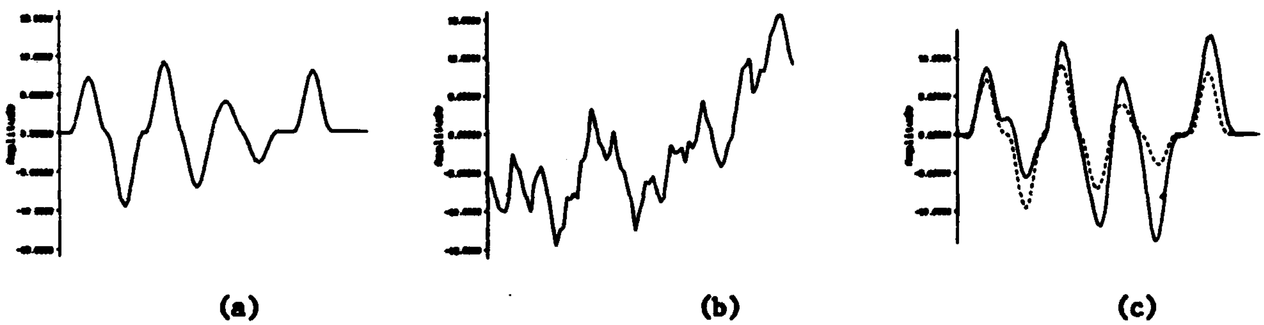


Figure 1. **Time-varying filter performance. (a) Underlying signal. (b) Signal plus noise. (c) Filtered waveform. All waveforms are 224 ms in duration.**

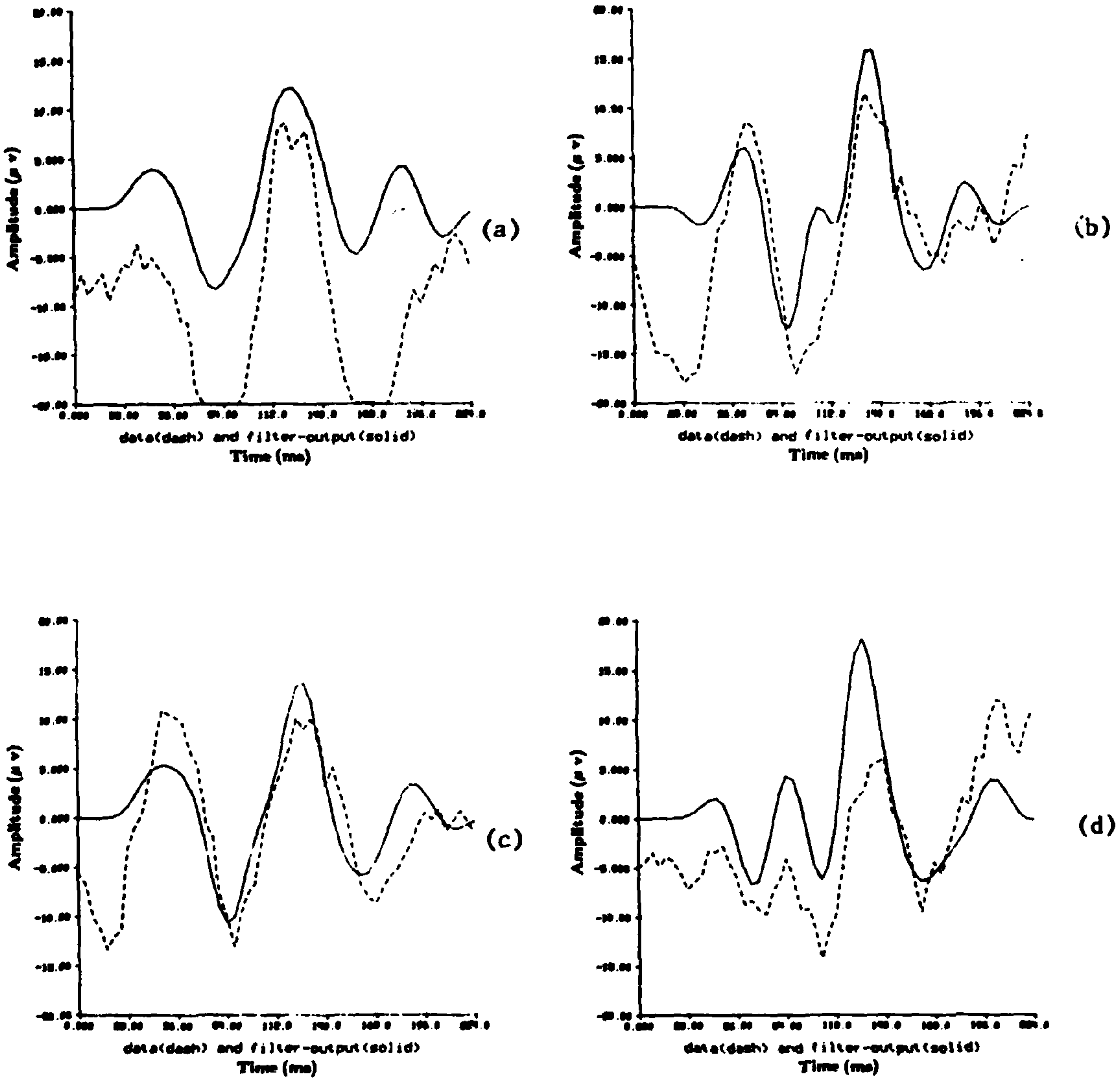


Figure 2. Evolved potential waveforms before (dashed) and after (solid) processing by the time-varying filter.

for quantitative measurements on single EP waveforms is being evaluated.

Effects of Noise on Latency Measurements of EP Components. Whenever repetitive measurements of evoked potentials are made it is found that there are significant variations in the amplitudes and latencies in the individual peaks or components that are present. This is a well known phenomenon and has frequently been discussed in the literature. Figure 3 shows a histogram of the latencies of peaks identified in 100 visual evoked potentials elicited by a flash stimulus for a subject with eyes closed. It is seen from the figure that the measured latencies are distributed around a mean value in a manner suggesting that a certain amount of randomness is associated with their occurrence.

There are two possible causes for the randomness associated with measurements of this kind. First, the measurements of EP's are made in the presence of the ongoing EEG which may significantly affect the shape of the observed waveform and alter both latency and amplitude estimates. Second the EP waveform itself may be varying from stimulus to stimulus. It is important for several reasons to be able to differentiate between these two effects. One reason is to better define the characteristics and parameters associated with the single EP. Another is to provide quantitative information for the design of improved signal processors for analysis and classification of single EPs. It is this latter reason that led to the research described here which is aimed at establishing the degree to which the presence of the ongoing EEG considered as an additive noise component can affect the latency measurements of components in the EP.

This problem was attacked theoretically and the resulting analytical expressions then checked empirically. The analysis was carried out as follows. It was assumed that an EP waveform could be approximated in the vicinity of a peak by a second order polynomial. The coefficients of the polynomial were determined by means of a least squares fitting of 5 points in the vicinity of a peak. The errors in estimating these parameters in the presence of noise were also determined theoretically. Once the parameters are known the peak location is readily found. The variance in the peak location estimate due to the presence of noise can be determined from the errors that occur in the parameter estimates. A simulation was performed to verify the theoretical results. Noise was added to a known waveshape and the peak location determined. This was repeated many times and the variance of the peak location determined. Tests were carried out utilizing both white noise and noise having the same covariance as measured EEG signals. It was found that white noise gave a slightly higher variance in the component latency than did the EEG noise. However, the check in both cases between the experimentally determined values and the theoretical values was quite close.

Latency variation or jitter due to noise depends primarily upon two parameters: the signal-to-noise-ratio and the radius of curvature of the peak. Narrow peaks and high signal-to-noise ratios reduce the latency jitter produced by additive noise. The following expression for the standard deviation of latency jitter was obtained.

$$\sigma = \frac{0.86 R W_n}{\sqrt{SNR}}$$

where R is the radius of curvature of the peak, W_n is the noise bandwidth and SNR is the signal-to-noise ratio computed as the square of the peak signal amplitude divided by the variance of the noise. This equation is valid for any sampling frequency equal to or greater than $2W_n$. A convenient pulse shape for computation purposes is a single lobe of a cosine wave. If the duration of this pulse is T seconds, then the radius of curvature is $(\frac{T}{\pi})^2$. To illustrate the nature of the results consider a case in which an EP component has a duration of 30 ms, the noise bandwidth is 25 Hz and the SNR is 0dB. The standard deviation of the latency jitter for this case would be 2 ms. Values determined from the LCA are typically 7 to 10 ms for measured EPs having similar parameters. This indicates that the variations in latency must be almost entirely due to variations in the latency of the components themselves and not due to the effects of additive noise. These results add considerable confidence to the design procedures for the improved signal processing techniques that make use of the random variations in signal latency.

PUBLICATIONS

J. I. Aunon, C. D. McGillem and D. G. Childers, "Signal Processing in Evoked Potential Research: Averaging and Modeling," *Critical Rev. In Bioengineering,* Vol. 5, July 1981, pp. 323-367.

D. G. Childers, J. I Aunon and C. D. McGillem, "Spectral Analysis and Extrapolation," *Critical Rev. in Bioengineering* Vol. 6, Sept. 1981, pp. 133-175.

K.B. Yu, C. D. McGillem, "Optimum Time Varying Filters for Transient Waveform Estimation," 19th Allerton Conf. on Com. Control, and Comp., Monticello,Ill., Sept. 30-Oct. 2, 1981.

C. D. McGillem, J. I. Aunon and D. G.Childers, "Signal Processing in Evoked Potential Research: Applications of Filtering and Pattern Recognition," *Critical Rev. in Bioengineering,* Vol. 7, Oct. 1981, pp. 225-265.

J. I. Aunon and C. D. McGillem, "On the Classification of Single Evoked Potentials Using a Quadratic Classifier," *Computer Prog. in Biomedicine,* Vol. 14-1, Feb. 1982.

C. D. McGillem, C. A. Pomalaza and J. I Aunon, "Preprocessing for Improved Classification of Evoked Potentials," *IEEE 3rd Annual Conf. of EMBS,* Sept 1981, Houston, TX.

PUBLICATIONS IN PREPARATION

(1) J. I. Aunon, C. D. McGillem, R. D. O'Donnell, "Comparison of Linear and Quadratic Classification of Event Related Potentials on the Basis of Their Exogenous or Endogenous Components," to be submitted to *Psychophysiology.*

(2) C. D. McGillem, J. I. Aunon and C. A. Pomalaza, "Improved Waveform Estimation Procedures for Event Related Potentials," to be submitted to *Psychophysiology.*

(3) C. D. McGillem, K. B. Yu and J. I. Aunon, "Effects of Ongoing EEG on Latency Measurements of Evoked Potentials," *IEEE Eng. in Medicine and Biology Conf.,* Sept. 1982.

(4) K. B. Yu and C. D. McGillem, "Matrix Processors for Estimating Nonstationary Random Waveforms," Conf. on Signal Recovery and Synthesis with Incomplete Information and Partial Constraints," Jan. 1983.

PROFESSIONAL PERSONNEL

Principal Investigators

Clare D. McGillem, Ph.D. Professor of Electrical Engineering
Jórge I. Aunon, D.Sc.. Associate Professor of Electrical Engineering

DEGREES AWARDED

Daniel L. Halliday, MSEE. Thesis: "Optimal and Suboptimal Feature Selection Algorithms," August 1981

Maryann Walsh, MSEE. Thesis: "Representation of Brain Evoked Potential Ensembles Using Optimized Complex Exponentials," August 1981.

Jon S. Williams, MSEE. Thesis: "Detection of Auditory and Visual Evoked Potentials in the Electroencephlagram Using a Maximum Signal to Noise Ratio Filter," August 1981.

John J. Westerkamp, MSEE. Thesis: "A Comparison of Conventional and Latency Corrected Averaging for the Processing of Event Related Brain Potentials," December 1981.